

CLOSED, OPEN AND MIXED QUEUEING NETWORKS WITH DIFFERENT DISTRIBUTION

PRASHANT CHAUHAN

Department of Mathematics, S.L. Education Institute, Moradabad, Uttar Pradesh, India

ABSTRACT

In this paper we will derive the equilibrium distribution of states of a model containing four different types of service centers and R different classes of customers. From this steady state distribution one can compute the moments of the queue sizes for different classes of customers at different service centers, the utilizations of the service centers, the “cycle time” or response time for different classes of customers, the “throughput” of different classes of customers, and other measures of system performance.

KEYWORDS: Laplace Ttransforms, Steady State Distribution, Queuing Networks $M/M/1$ & $M/G/\infty$ Modal

INTRODUCTION

These results unify and extend a number of separate results on networks of queues. The general model can have four types of service centers. Three of those types allow different service time distributions with rational **Laplace** transforms for different classes of customers. The model allows different classes of customers to have different arrival rates and different routing probabilities. For open networks some very simple formulas give the marginal distribution of customers at the service centers of the network.

The analysis is motivated by the desire to model computer systems. Type one service centers (FCFS scheduling) seem appropriate models of secondary storage input/output devices. Type two service centers (processor sharing scheduling) can be an appropriate model for central processing units. Type three service centers, (no queueing) are appropriate models for, terminals and for routing delays in the network.

The requirement that a service time distribution have a rational Laplace Transform is not very restrictive. Exponential, hyperexponential and hypoexponential distributions all have rational **Laplace** Transforms. Cox [12] has shown that any such distribution can be represented by a network of exponential stages of the form illustrated in Fig. 1. For convenience, we have eliminated the case in which there is a non-zero probability of a zero length service time.

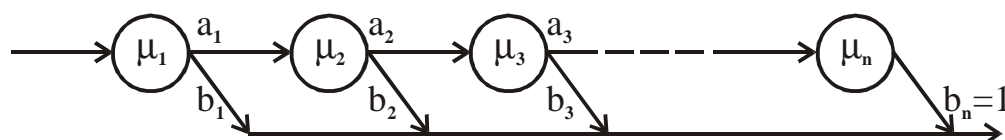


Figure1

In this figure, b_i is the probability that the customer leaves after the i^{th} stage and $a_i (= 1 - b_i)$ is the probability that the customer goes to the next stage. Given that a customer reaches the i^{th} stage the service time in this stage has a

negative exponential distribution with mean $1/\mu_i$. Since the service time distribution for a stage is exponential, when describing the state of the network of service stations it is not necessary to know the exact amount of service a customer has received at a service center; the stage of service is sufficient.

In [4], Chandy terms these the global balance equations. He defines another type of balance equation which he calls the local balance equations. Informally, a local balance equation equates the rate of flow into a state by a customer entering a stage of service to the flow out of that state due to a customer leaving that stage of service. We associate a customer with a stage of service in the following ways. If the customer is in service at a service center, then he is in one of the stages of his service time distribution at that service center. If the customer is queued at a service center, then he is in the stage of his service time distribution he will enter when next given service.

The local balance equations are sufficient conditions for global balance, but they are not necessary. Local balance requires, that each term on the right-hand side of a global balance equation be equal to a particular subset of terms on the left-hand side of that global balance equation.

To illustrate the concept of local balance we consider the relatively simple network model in Figure. 2.

This is a closed network with two classes of customers (which we refer to as class 1 and class 2). There are N_1 class 1 customers and N_2 class 2 customers in the networks. All service times are exponentially distributed and

$\frac{1}{\mu_{ir}} (i = 1, 2, r = 1, 2)$ is the mean service time for a class r customer at service center i .

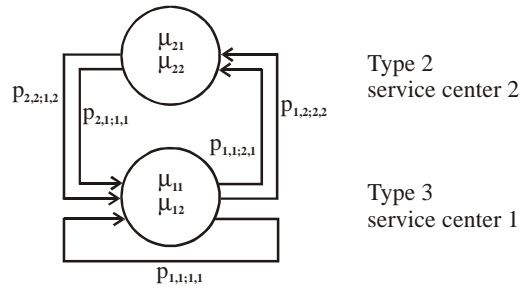


Figure 2

In this example $P_{1,2;2,2} = P_{2,2;1,2} = P_{2,1;1,1} = 1$, $P_{1,1;1,1} + P_{1,1;2,1} = 1$.

Let $n_{i,r}$ be the number of class r customers at service center i . For convenience we write the global and local balance equations only for the states in which $n_{ir} > 0$, $i = 1, 2, r = 1, 2$.

Global Balance Equation:

$$\begin{aligned}
 &P(n_{11} - 1, n_{12}, n_{21} + 1, n_{22}) \left(\frac{n_{21} + 1}{n_{21} + n_{22} + 1} \right) \mu_{21} \\
 &+ P(n_{11} + 1, n_{12}, n_{21} - 1, n_{22}) (n_{11} + 1) \mu_{11} P_{1,1;2,1} \\
 &+ P(n_{11}, n_{12}, n_{21}, n_{22}) n_{11} \mu_{11} P_{1,1;1,1} + P(n_{11}, n_{12} + 1, n_{21}, n_{22} - 1) (n_{12} + 1) \mu_{12}
 \end{aligned}$$

$$\begin{aligned}
& +P(n_{11}, n_{12}-1, n_{21}, n_{22}+1) \left(\frac{n_{22}+1}{n_{21}+n_{22}+1} \right) \mu_{22} \\
& = P(n_{11}, n_{12}, n_{21}, n_{22}) \left[n_{11} \mu_{11} + n_{12} \mu_{12} + \frac{n_{21}}{n_{21}+n_{22}} \mu_{21} + \frac{n_{22}}{n_{21}+n_{22}} \mu_{22} \right]
\end{aligned}$$

Local Balance Equations:

$$P(n_{11}-1, n_{12}, n_{21}+1, n_{22}) \left(\frac{n_{21}+1}{n_{21}+n_{22}+1} \right) \mu_{21} \quad (1.1)$$

$$+P(n_{11}, n_{12}, n_{21}, n_{22}) n_{11} \mu_{11} P_{1,1;1,1} = P(n_{11}, n_{12}, n_{21}, n_{22}) n_{11} \mu_{11}$$

$$P(n_{11}, n_{12}-1, n_{21}, n_{22}+1) \left(\frac{n_{22}+1}{n_{21}+n_{22}+1} \right) \mu_{22} = P(n_{11}, n_{12}, n_{21}, n_{22}) n_{12} \mu_{12} \quad (1.2)$$

$$P(n_{11}+1, n_{12}, n_{21}-1, n_{22}) (n_{11}+1) \mu_{11} P_{1,1;2,1} = P(n_{11}, n_{12}, n_{21}, n_{22}) \left(\frac{n_{21}}{n_{21}+n_{22}} \right) \mu_{21} \quad (2.1)$$

$$P(n_{11}, n_{12}+1, n_{21}, n_{22}-1) (n_{12}+1) \mu_{12} = P(n_{11}, n_{12}, n_{21}, n_{22}) \left(\frac{n_{22}}{n_{21}+n_{22}} \right) \mu_{22} \quad (2.2)$$

Since all the service time distributions in this, example are exponential the current stage of service of a customer is uniquely defined by the customer's class and the current service center. Local balance equation (i,r) for $i=1,2$, $r=1,2$ equates the rate of flow out of state $(n_{11}, n_{12}, n_{21}, n_{22})$ due to a class r customer leaving service center i with the rate of flow into state $(n_{11}, n_{12}, n_{21}, n_{22})$ due to a class r customer entering service center i .

As in this example it is generally true that each global balance equation is the sum of a sub-set of the local balance equations. Thus a solution for the local balance equations is automatically a solution to the global balance equations. In many cases the local balance equations are inconsistent and therefore have no solution. For example if there is FCFS scheduling at a service center and different classes of customer have different service time distributions the local balance equations are inconsistent.

The value of the local balance concept is that (1) it leads to a simpler and more organized search for solutions for equilibrium state probabilities and (2) it works for a large number of cases (in fact for virtually all of the closed form solutions known for general classes of networks of queues--although not many interesting cases have known solutions).

Before presenting the solution to the class of networks described, we define a set of terms that appear in the solution.

For a customer of class r , let $\{e_{ir}, 1 \leq i \leq N\}$ be a solution to the following set of equations.

$$\sum_{1 \leq i \leq N} e_{ir} P_{i,r;j,s} + d_{js} = e_{js}, \quad 1 \leq j \leq N$$

The value of d_{js} is determined by the arrival process of customers of class s to service center j . If there are no such arrivals from outside the system, then $d_{js} = 0$. If there are such arrivals then $d_{js} = q_{js}$. In a closed system there are no arrivals to any center and all the d_{js} are zero. In this case e_{ir} is the relative frequency of visits to service center i by customers of class r .

Note that a system may be “open” with respect to some classes of customers and “closed” with respect to other classes of customers. Our solution applies to this class of system.

One further definition is required. If at the i^{th} service center the r^{th} class of customers has a service time distribution that is represented as a network of stages then this is represented as illustrated in figure 3.

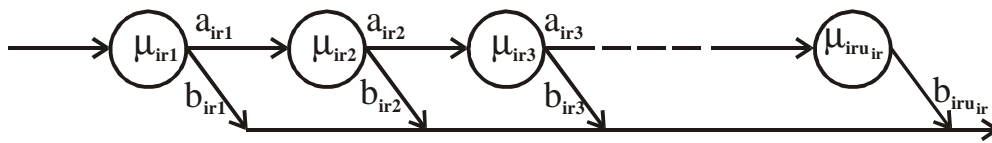


Figure 3

The first subscript on a , b and μ denotes the service center. The second subscript denotes the class of customer and the third subscript denotes the stage.

$$\text{Let } A_{ir\ell} = \prod_{j=1}^{\ell} a_{irj}$$

Theorem

Given a network of service stations which is open, closed or mixed in which each service center is of type 1, 2, 3 or 4. Then the equilibrium state probabilities are given by

$$P(S = x_1, x_2, \dots, x_N) = Cd(S) f_1(x_1) f_2(x_2) \dots f_N(x_N)$$

where C is a normalizing constant chosen to make the equilibrium state probabilities sum to 1, $d(S)$ is a function of the total number of customers in system and each f_i is a function that depends on the type of service center i .

If service center i is of type 1 then

$$f_i(x_i) = \prod_{j=1}^{n_i} \left[\frac{1}{\mu_i(j)} e^{ix_{ij}} \right]$$

If service center i is of type 2 then

$$f_i(x_i) = n_i! \prod_{r=1}^R \prod_{\ell=1}^{u_{ir}} \left\{ \left[\frac{e_{ir} A_{ir\ell}}{u_{ir\ell}} \right]^{m_{ir\ell}} \frac{1}{m_{ik\ell}!} \right\}$$

If service center i is of type 3 then

$$f_i(x_i) = \prod_{r=1}^R \prod_{\ell=1}^{u_{ir}} \left\{ \left[\frac{e_{ir} A_{ir\ell}}{u_{ir\ell}} \right]^{m_{ir\ell}} \frac{1}{m_{ir\ell}!} \right\}$$

If service center i is of type 4 then

$$f_i(x_i) = \prod_{j=1}^{n_i} \left[e_{ir_j} A_{ir_j} \frac{1}{m_j \mu_{ir_j m_j}} \right]$$

If the arrivals to the system depend on the total number of customers in the system, $M(S)$, and the arrivals are of class r and for center i according to fixed probabilities P_{ir} then

$$d(S) = \prod_{i=0}^{M(S)-1} \lambda(i)$$

If we have the second type of state dependent arrival process then

$$d(S) = \prod_{j=1}^m \prod_{i=0}^{M(E_j)-1} \lambda_j(i)$$

If the network is closed then $d(S) = 1$.

The theorem is proved by checking that the local balance equations are satisfied. In every case for which these results apply the local balance equations reduce to the defining equations for the $\{e_{ir}\}$.

Simplification of Results

The solution presented for the equilibrium state probabilities deals with system states that are more detailed than is usually required. The more detailed states are necessary to derive the equilibrium state probabilities. Now we define the system state as the number of each class of customer in each service center. More formally state S of the system is given by (y_1, y_2, \dots, y_N) where $y_i = (n_{i1}, n_{i2}, \dots, n_{ir})$ and n_{ir} is the number of customers of class r in service center i . Let n_i

be the total number of customers at service center i and let $\frac{1}{\mu_{ir}}$ be the mean service time of a class r customer at service center i . Then the equilibrium state probabilities are given by

$$P(S = (y_1, y_2, \dots, y_N)) = Cd(S) g_1(y_1) g_2(y_2) \dots g_N(y_N)$$

where

if service center i is of type 1 then

$$g_i(y_i) = n_i! \left\{ \prod_{r=1}^R \frac{1}{n_{ir}!} [e_{ir}]^{n_{ir}} \right\} \prod_{j=1}^{n_i} \frac{1}{\mu_i(j)}$$

if service center i is of type 2 or 4 then

$$g_i(y_i) = n_i! \prod_{r=1}^R \frac{1}{n_{ir}!} \left[\frac{e_{ir}}{\mu_{ir}} \right]^{n_{ir}}$$

if service center i is of type 3 then

$$g_i(y_i) = \prod_{r=1}^R \frac{1}{n_{ir}!} \left[\frac{e_{ir}}{\mu_{ir}} \right]^{n_{ir}}$$

In each case the expression for $g_i(y_i)$ is derived by summing $f_i(x_i)$ over all x_i with $n_{i1}, n_{i2}, \dots, n_{ik}$ fixed. That this is the correct definition of the g_i follows from the product form of the solution given

The evaluation of the normalizing constant requires summing the given expression for the equilibrium state probabilities over all feasible states. In the next section we show a closed form solution for C for an open network in which $\mu_i(m) = \mu_i$ for all m if service center i is of type one.

Open Systems

For open systems it is possible to obtain a closed form solution for the normalization constant when the arrival process is of the first type and $\lambda(M(S)) = \lambda = \text{constant}$. Since the system is open any number of customers is feasible at a service center.

Therefore

$$c^{-1} = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \dots \sum_{n_N=0}^{\infty} \left(\prod_{i=1}^N \lambda^{n_i} h_i(n_i) \right)$$

$$\text{or} \quad c^{-1} = \left(\sum_{n_1=0}^{\infty} \lambda^{n_1} h_1(n_1) \right) \left(\sum_{n_2=0}^{\infty} \lambda^{n_2} h_2(n_2) \right) \dots \left(\sum_{n_N=0}^{\infty} \lambda^{n_N} h_N(n_N) \right)$$

Also,

$$\sum_{n_i=0}^{\infty} h_i(n_i) = \left(1 - \sum_{r \in R_i} \lambda \frac{e_{ir}}{\mu_i} \right)^{-1} \quad \text{if service center } i \text{ is type 1 and } \mu_i(n_i) = \mu_i$$

$$= \left(1 - \sum_{r \in R_i} \lambda \frac{e_{ir}}{\mu_{ir}} \right)^{-1} \quad \text{if service center } i \text{ is type 2 or 4}$$

$$e^{\sum_{r \in R_i} \lambda \frac{e_{ir}}{\mu_{ir}}} \quad \text{if service center } i \text{ is type 3}$$

Marginal Distribution at a Service Center in an Open System

Let $P_i(n_i)$ be the equilibrium probability that there are n_i customers at service center i .

$$P_i(n_i) = C \lambda^{n_i} h_i(n_i) \prod_{\substack{j=1 \\ j \neq i}}^N \left(\sum_{n_j=0}^{\infty} \lambda^{n_j} h_j(n_j) \right)$$

Using the expression for C , we reduce this to

$$P_i(n_i) = \frac{\lambda^{n_i} h_i(n_i)}{\sum_{m=0}^{\infty} \lambda^m h_i(m)}$$

Let
$$\rho_i = \sum_{r \in R_i} \lambda \frac{e_{ir}}{\mu_i} \quad \text{if service center } i \text{ is type 1.}$$

$$\rho_i = \sum_{r \in R_i} \lambda \frac{e_{ir}}{\mu_{ir}} \quad \text{if service center } i \text{ is type 2, 3 or 4.}$$

Then
$$P_i(n_i) = (1 - \rho_i) \rho_i^{n_i} \quad \text{if service center } i \text{ is type 1, 2 or 4.}$$

$$= e^{-\rho_i} \frac{\rho_i^{n_i}}{n_i!} \quad \text{if service center } i \text{ is type 3.}$$

These results provide a convenient way of examining the equilibrium distribution at a service center, For type 1, 2 or 4 service stations the marginal distribution is the same as the distribution of the number of customers in an $M/M/1$ queue with a suitably chosen utilization, ρ_i . For the equilibrium solution to exist each ρ_i is required to be less than 1.

The marginal distribution for a type 3 service center is the same as the equilibrium distribution for the number of customers for an $M/G/\infty$ system with $\rho_i = \frac{\lambda}{\mu}$. This certainly appears to be reasonable since for an open properties of network models that satisfy local balance.

All of the network models that we have treated in this paper can be shown to be equivalent to models in which all classes of customers have the same exponential service time distribution at a given service center. Thus an exponential service time distribution with mean $\frac{1}{\mu_i}$ may be associated with the i^{th} service center and all classes of customers have this service time distribution at the i^{th} service center. This fact suggests the conjecture that a necessary condition for local balance to be satisfied for a given model is that there exist an equivalent model in which different classes of customers may have different transition probabilities but all classes of customers have the same exponential service time distribution at a given service center.

The method of making these transformations to the model is straight-forward and will be illustrated by example rather than a formal description of the general case.

Consider a general service time distribution represented by a network of stages as in Figure 1. Let this represent the service time distribution for a customer in class r in service center i . We introduce n new customer classes denoted by r_1, r_2, \dots, r_n which correspond to the stages in this network and delete customer class r .

The service time of a class r_ℓ customer will be exponential with mean $\frac{1}{\mu_\ell} (1 \leq \ell \leq n)$. The transition probabilities for a class r_ℓ customer are defined as:

$$p_{i,r_\ell;j,s} = \ell P_{i,r;j,s}$$

$$p_{i,r_\ell;i,r_{\ell+1}} = a_\ell \quad 1 \leq \ell < n$$

To take care of the transitions into class r in the original model we require that all transitions into state r be redefined as transitions into state r_1 . These transition probabilities are defined as:

$$P_{j,s;i,r_1} = P_{j,s;i,r}, \quad j, s$$

With this transformation of the model a customer will have the same distribution of total time at a service center and will have the same transition probabilities from service center to service center.

After performing this transformation for each customer class with a general service time distribution we have a model in which all service time distributions are exponential. Suppose that $\frac{1}{\mu_{i,r}}$ is the mean service time of a class r customer at service center i . Let

$$\mu_i = \max_r [M_{i,r}]$$

We redefine the mean service time for each class of customers at service center i to be $\frac{1}{\mu_i}$. Now we redefine the transition probabilities out of service center i ,

$$\text{Let } P_r = 1 - \frac{\mu_{i,r}}{\mu_i}$$

Then define

$$P'_{i,r;i,r} = P_r + (1 - p_r) P_{i,r;i,r}$$

$$P'_{i,r;j,s} = (1 - p_r) P_{i,r;j,s}$$

The effect of these new transition probabilities is to cause a class r customer to be fed back (or to revisit) service center i a random number of times. Each time the class r customer enters service center i his service time is exponentially distributed with mean $\frac{1}{\mu_i}$. The number of visits the class r customer makes to service center i (between transitions in the

original model) is geometrically distributed with mean $\frac{1}{1-p_r}$. It is easily shown that the total service time of the class r customer at service center i is exponentially distributed with mean $\left(\frac{1}{1-p_r}\right)\frac{1}{\mu_i} = \frac{1}{\mu_{i,r}}$ [8]. Therefore we have not changed the total service time distribution for this customer at service center i .

After completing these transformations throughout the model we have an equivalent model with the desired characteristics.

CONCLUSIONS

The transformations that we have made to the original model preserved the original distributions of service time that a customer requires at a service center. However a customer does not spend that time on the server in one contiguous interval. We required a customer to make extra transitions in which he leaves and reenters the service center. It is clear that with type 2, 3 or 4 service centers this does not affect a customers service. For type 1 service centers the transformed model would not be equivalent to the original model since a customer who leaves the service center and reenters will now be at the end of the queue. Of course we have from the beginning required that at type 1 service centers all customers have the same exponential service time distribution so that such a service center does not require any modification.

REFERENCES

1. Baskett, F. The dependence of computer system queues upon processing time distribution and central processor scheduling, Proc. of the ACM SIGOPS Third Symposium on Operating System Principles, Stanford University, (October 1971), 109-113.
2. Baskett, F. and Palacios, F. G. Processor sharing in a central server queueing model of multiprogramming with applications, Proc. of the Sixth Annual Princeton Conference on Information Sciences and Systems, Princeton University, (March 1972).
3. Buzen, J. Queueing Network Models of Multiprogramming, Ph.D. Thesis, Division of Engineering and Applied Science, Harvard University, Cambridge, Mass., 1971.
4. Chandy, K.M. The analysis and solutions for general queueing networks, Proc. of the sixth Annual Princeton Conference on Information Sciences and Systems, Princeton University, (March 1972).
5. Chandy, K. M. Exponential and processor-sharing queueing network models for computer systems, IEEE Trans. on Computers, (to appear).
6. Chandy, K. M., Keller, T. W., and Browne, J. C. Design automation and queueing networks: an interactive system for the evaluation of computer queueing models, Proc. of the **Ninth** Annual Design Automation Workshop, Dallas, June, **1972**.
7. Cox, D.R. A use of complex probabilities in the theory of stochastic processes, Proc. Carob. Phil, Sox 51 (1955), 313-319.
8. Feller, W. An Introduction to Probability Theory and its Applications, Vol. 1, 3rd ed., Wiley, New York, 1968.

9. Ferdinand. A.E. An analysis of the machine interference model, IBM Systems J., 10, 2 (1971), pp. 120-142.
10. Gordon, W. J. and Newell, G, F. Closed queueing systems with exponential servers, Opns. Res, 15 (1967), 254-26.
11. Jackson, J. R. Jobshop-like queueing systems, Man. Sci. 10, 1 (Oct. 1963), pp. 131–142.
12. Moore, C.G., III. Network Models for Large-Scale Time-Sharing Systems, Technical Report No. 71–1, Department of Industrial Engineering, The University of Michigan, Ann Arbor, Michigan, (April 1971).
13. Palacios, F.G. An analytic model of a multiprogramming system including a job mix, Report TR-4, Department of Computer Sciences, University of Texas at Austin, June, 1972.
14. Posner, M. and Bernholtz, B. Closed finite queueing networks with time lags and with several classes of units, Opns. Res. 16 (1968), PP. 977-985.
15. Sakata, M., Noguchi, S., and Oizumi, J., Analysis of a processor-shared queueing model for time-sharing systems, proc. of the Second Hawaii International Conference on System Sciences, January 1969.
16. Whittle, P. Nonlinear migration processes. Proc. 36th Session of the International Statistical Institute, pp. 642-647.
17. Whittle, P. Equilibrium distributions for an open migration process. J. Appl. Prob. 5, pp. 567-571.